

UCLA

UCLA Previously Published Works

Title

Partial duplication of the APBA2 gene in chromosome 15q13 corresponds to duplicon structures.

Permalink

<https://escholarship.org/uc/item/3x33g94s>

Journal

BMC genomics, 4(1)

ISSN

1471-2164

Authors

Sutcliffe, James S
Han, Michael K
Amin, Taneem
et al.

Publication Date

2003-04-01

DOI

10.1186/1471-2164-4-15

Peer reviewed

Research article

Open Access

Partial duplication of the *APBA2* gene in chromosome 15q13 corresponds to duplicon structures

James S Sutcliffe*, Michael K Han, Taneem Amin, Robert A Kesterson and Erika L Nurmi

Address: Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN, USA

Email: James S Sutcliffe* - james.s.sutcliffe@vanderbilt.edu; Michael K Han - Michael_K_Han@hotmail.com; Taneem Amin - taneem.amin@vanderbilt.edu; Robert A Kesterson - robert.kesterson@vanderbilt.edu; Erika L Nurmi - erika.nurmi@vanderbilt.edu

* Corresponding author

Published: 29 April 2003

Received: 22 February 2003

BMC Genomics 2003, 4:15

Accepted: 29 April 2003

This article is available from: <http://www.biomedcentral.com/1471-2164/4/15>

© 2003 Sutcliffe et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Chromosomal abnormalities affecting human chromosome 15q11-q13 underlie multiple genomic disorders caused by deletion, duplication and triplication of intervals in this region. These events are mediated by highly homologous segments of DNA, or duplicons, that facilitate mispairing and unequal cross-over in meiosis. The gene encoding an amyloid precursor protein-binding protein (*APBA2*) was previously mapped to the distal portion of the interval commonly deleted in Prader-Willi and Angelman syndromes and duplicated in cases of autism.

Results: We show that this gene actually maps to a more telomeric location and is partially duplicated within the broader region. Two highly homologous copies of an interval containing a large 5' exon and downstream sequence are located ~5 Mb distal to the intact locus. The duplicated copies, containing the first coding exon of *APBA2*, can be distinguished by single nucleotide sequence differences and are transcriptionally inactive. Adjacent to *APBA2* maps a gene termed *KIAA0574*. The protein encoded by this gene is weakly homologous to a protein termed X123 that in turn maps adjacent to *APBA1* on 9q21.12; *APBA1* is highly homologous to *APBA2* in the C-terminal region and is distinguished from *APBA2* by the N-terminal region encoded by this duplicated exon.

Conclusion: The duplication of *APBA2* sequences in this region adds to a complex picture of different low copy repeats present across this region and elsewhere on the chromosome.

Background

Human chromosome 15q11-q13 is associated with genomic disorders involving deletions, duplications and triplications of different intervals in this region. The formation of these cytogenetic abnormalities is mediated by low-copy repeated sequences, or duplicons [1]. High sequence homology between copies presumably facilitates mispairing and unequal recombination events in meiosis. Interstitial deletion of 15q11-q13 results in Prader-Willi

syndrome (PWS; MIM # 176270) or Angelman syndrome (AS; MIM # 105830), when deletions occur on the paternally-derived or maternally-derived homolog, respectively (see Figure 1; reviewed in refs. [2-4]). Interstitial duplication of the same interval has been found in cases of autism, and data suggest a higher risk for autism with maternal compared to paternal duplications [5-7]. Chromosome 15 is the most frequently involved chromosome in cases of autism with supernumerary marker

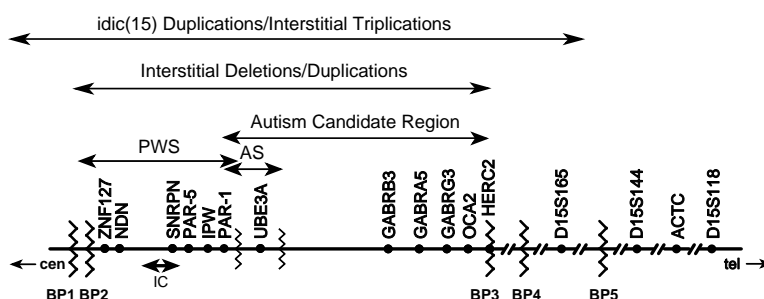


Figure 1

Schematic representation of 15q11-q13. The chromosome 15q11-q13 is depicted, with breakpoints utilized in PWS/AS deletions and autism-associated duplications and triplications represented by larger jagged lines. Smaller jagged lines identify rare chromosomal breakpoints which delimit the PWS and AS critical regions. Centromeric (cen) and telomeric (tel) orientation is shown and breakpoints are numbered from BP1-BP5. Arrows over the map show the intervals deleted in PWS/AS and duplicated in duplications or triplications affecting 15q11-q13. Gene and framework marker positions are indicated over the map. The most narrow autism candidate region is based on the smallest autism-associated maternally-biased duplications.

chromosomes. 15q-derived markers are inverted, duplicated and pseudodicentric structures and are known as inv dup(15) or idic(15). Small markers which do not contain the PWS/AS deletion interval are associated with a normal phenotype [8]. Larger idic(15) duplications are typically associated with a more severe autistic phenotype than that seen in cases of interstitial duplication [5-7,9-12]. The increased severity may be due to the presence of two additional copies of this region, however these duplications also extend significantly farther telomerically and therefore contain more duplicated genes. Idic(15) markers also vary in the location of the distal breakpoint, depending on which duplcon sequence mediated the rearrangement. Interstitial triplications have also been described, and these tend to utilize variable distal breakpoints common to idic(15) chromosomes [13]. Triplications are associated with a phenotype analogous to that for idic(15) when maternally-derived [7,12,13].

Several functional duplcons have been described, and these may be divided into two classes. There are two proximal breakpoints (BP1 and BP2) that define two classes of PWS/AS deletion [14]. The common distal breakpoint is termed BP3 and has a bipartite structure [15,16]. The BP2 and BP3 duplcons span several hundred kb and contain at least seven expressed sequences. One prominent example is the *HERC2* (Hect domain and Rcc1 domain protein 2) gene [17]. Approximately 11 *HERC2* copies have been identified on chromosomes 15 and 16, although most map to 15q11-q13. Another class of 15q duplcons has been termed LCR15s, for low copy repeats [18,19]. These share some sequences in common with the large *HERC2* duplcons. LCR15s extend only ~15 kb individually, but

are often present in clusters extending to 30-60 kb. This class of duplcon has been proposed to help mediate rearrangements in idic(15) markers and larger interstitial duplications and triplications [19].

The gene encoding the amyloid precursor protein-binding protein A2 (*APBA2*) was previously mapped to chromosome 15 [20] and localized to the common PWS/AS deletion region centromeric to BP3 and distal to a cluster of GABA_A receptor subunit genes [21,22] (see Figure 1). This location made *APBA2* an attractive positional candidate gene for involvement in duplication-associated and inherited autism. On this basis we sought to characterize *APBA2* in terms of structure and expression. *APBA2* has also been termed Mint2, for Munc-interacting protein 2, or X-11β. Mint2 functions as a neuronal adaptor protein that binds to the Munc18-1 protein, which is essential for synaptic vesicle exocytosis [23]. Biederer and Südhof have shown that a multiprotein complex containing Mint1 (*APBA1*) binds to the cytoplasmic tail of neurexins, which play a critical role in synaptic vesicle exocytosis. Mint1 and Mint2 can bind directly to neurexins via a PDZ domain-mediated interaction. Thus the Mint proteins serve to anchor Munc-18 to the neurexins and thereby facilitate exocytosis. We show here that *APBA2* actually maps to a more distal location, outside the narrow common interstitial deletion/duplication interval but within the larger idic(15) duplications. We found that a large 5' and the first coding exon of *APBA2* is duplicated, and these transcriptionally inactive copies are located at more telomeric sites within the region. The nature of the duplication may have implications for the evolution of *APBA2*/Mint2 from its other family members.

Results

We sought to characterize the structure and expression of the *APBA2* gene, as a prelude to analysis of this gene as a possible autism candidate. We obtained a bacterial artificial chromosome (BAC) clone (686I6) containing genomic sequence of the *APBA2* transcriptional unit. We also obtained yeast artificial chromosome (YAC) clones covering the interval between the 3' end of *GABRG3* and *OCA2* (962D11; contig WC-153) and to which an *APBA2* EST (W1-14097) and nearby markers had been mapped (764C6; contig WC-699). STS content mapping demonstrated that *APBA2* sequences and other markers are present within the BAC 686I6 and YAC 764C4 but absent from the YAC clone covering the *GABRG3-OCA2* interval (Figure 2b). Thus, the previous map position appeared to be incorrect. Based on STS marker content in the 686I6 clone, *APBA2* maps to a more distal location in 15q13. The revised map position places *APBA2* outside of the most narrow interstitial duplication interval and the area most supported by linkage data but within the larger idi(15) duplications. Current genomic sequence assemblies now agree with this experimentally-determined data.

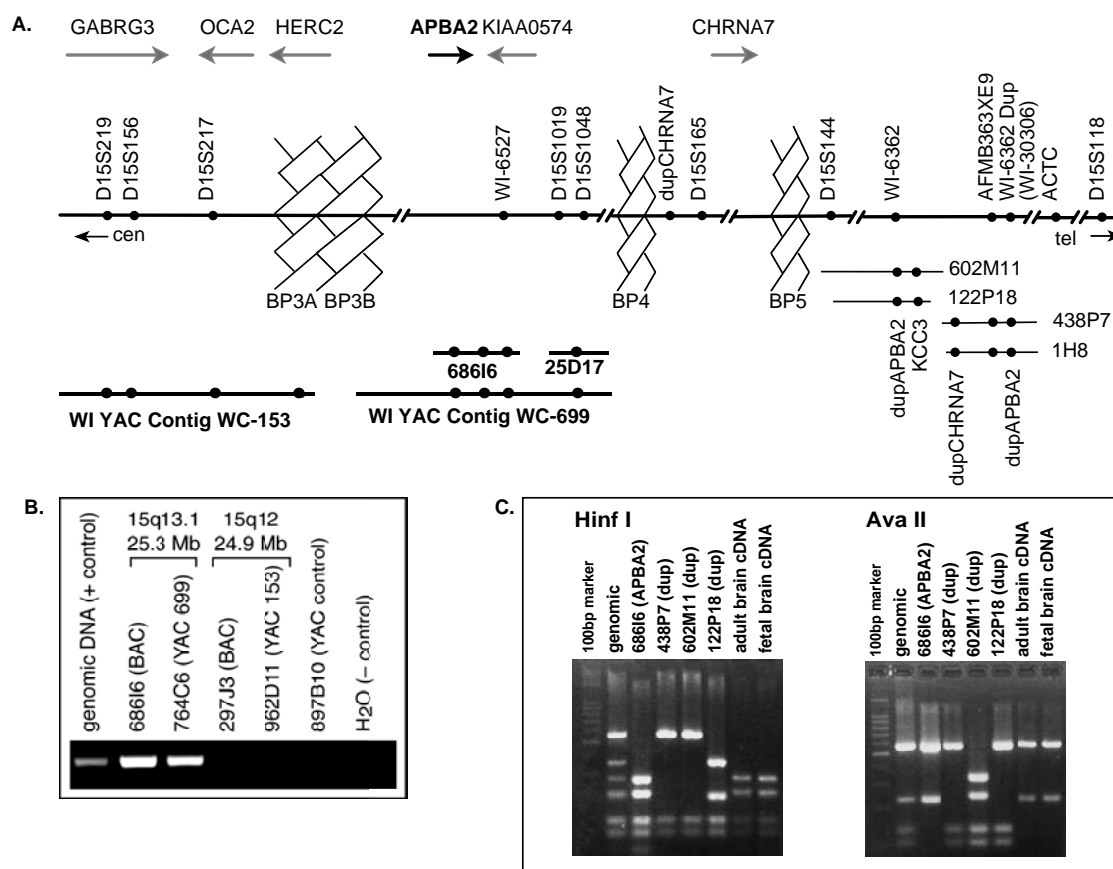
To facilitate direct screening of *APBA2*, we determined the gene structure and developed exon-specific PCR assays. We characterized genomic structure using a combination of existing genomic sequence and direct sequencing from BAC clone template. *APBA2* is encoded by 14 exons; 12 exons are coding and 2 exons contain 5' untranslated sequence. All exons are now accurately predicted in the NCBI (Genbank: NT_010363) and Celera assemblies, while the Ensembl assembly lacks the putative first exon. A sequence containing this exon has been deposited in GenBank (accession: AY228760). Junctions for most exons were confirmed by direct sequencing from RP-11 BAC 330K3 template, which also contains the *APBA2* transcriptional unit. Table 1 itemizes intron-exon boundary information. Primer sequences and conditions for PCR amplification of coding exons (3–14) are presented in Table 2. *APBA2* is transcribed towards the telomere.

In the course of analyzing the structure of *APBA2*, we noted that a large 5' exon containing the initiating methionine codon was present not only in the 686I6 BAC but also highly homologous to sequences from other BAC clones with non-overlapping STS content. *In silico* STS content mapping revealed two distinct pairs of BAC clones containing the partial duplications (Figure 2a). One of these (RP-11 clones 602M11 and 122P18) also contains a partial duplication of a pseudogene containing the 3' end of the neuronal nicotinic receptor $\alpha 7$ subunit (*CHRNA7*) gene [24]. This pseudogene locus maps to a site telomeric to *APBA2* in 15q13. The second group of BACs (RP-11 438P7 and 1 H8) contained the partial *APBA2* duplication and the K⁺Cl⁻ cotransporter (*SLC12A6*

or *KCC3*) gene. The duplications include the entire 1-kb exon and ~5-kb of downstream sequence, however this sequence is interrupted in the intact locus by an ~9.7-kb non-duplicated interval. The first block of homology includes exon 3, ~100 bp of sequence 5' and 1,045 bp 3' to the exon. Following the homology gap, there is an additional 3,945 bp of duplicated sequence. Furthermore, an additional 15 kb region in both sets of BACs containing partial *APBA2* duplications is highly homologous to a region approximately 25 kb upstream of the presumptive first exon of *APBA2*. Current sequence assemblies show the correct map location for the intact locus and show the map location of these dup*APBA2* sequences at ~5 Mb telomeric.

BLAST analysis of BACS containing the partial *APBA2* duplications revealed duplcon-like sequences, based on significant homology to a large number of BACs from known duplcons. The largest number of similarities detected (>30) corresponds to an interval of ~15 kb located ~15 kb centromeric to the *APBA2* duplications. However, there is highly significant homology over ~100-kb of sequence from assembly contig NT_035325 mapping to 15q26, where LCR-15 duplcon sequences have been reported [18]. These sequences appear to correspond to the LCR15 class of duplcons, based on sequence content. BLAST analysis of sequence at the intact locus revealed a similar low copy repeat or duplcon-type sequence of ~1.3-kb. More than 80 copies of this sequence are present at sites across the genome on every chromosome, and with an apparent clustering at telomere locations. This sequence, while short, does not correspond to any of the known chromosome 15q11-q13 duplcon sequences, and therefore could represent a novel class of such low copy repeats. This sequence has been deposited in GenBank (AY237156). It is worth noting that the presence of multiple duplcons and repeated sequences has significantly complicated genomic sequence assembly for this region.

Since the partial duplications of *APBA2* contained the first coding exon and the transcription start site(s) for *APBA2* have not been characterized, we sought to determine whether these copies were transcriptionally active. We developed restriction fragment length polymorphism (RFLP) assays to detect sequence differences discriminating the intact *APBA2* locus from the partial duplications. These assays were used following PCR from genomic or BAC DNA and cDNA from adult and fetal human brain (Figure 2c). Distinct "fingerprint" patterns are apparent for the intact locus and the duplications. The sequence differences used in Figure 2c that distinguish the duplications from the intact locus do not distinguish the duplications from each other. This distinction was made based on *in silico* analysis of numerous sequence differences, clone-sequence assembly and clone-marker/gene

**Figure 2**

Physical mapping of *APBA2* and PCR-RFLP analysis of *APBA2* duplications. A. An expanded view of the 15q12-q14 region containing distal duplication breakpoints is shown. Breakpoints BP3 and BP5 are depicted as jagged, hatched structures; BP4 is not thus shown but its position is indicated by an arrow over the map. STS marker positions are indicated on the map by closed circles, and arrows over the map indicate transcriptional orientation for genes in this region. YAC and BAC clone locations are shown below the map; YAC 764C6 is located in contig WC-153 and 962D11 is in WC-699. Closed circles within clones reflects the inclusion and position of specific STS markers shown above. B. PCR amplification of an *APBA2*-specific STS using genomic, YAC and BAC templates is shown. Position of clone pairs relative to cytogenetic banding and public sequence assemblies is indicated. Product consistent with genomic template is present only in clones from the region just telomeric to BP3. C. PCR from genomic or BAC clone template, followed by restriction with either *Hinf*I or *Ava*II is shown. Unique banding patterns, corresponding to single nucleotide differences between copies, may be observed for the intact locus and each of the two duplication copies. RT-PCR using adult or fetal cDNA reveals the pattern corresponding to the intact locus only.

content relationships. The pattern in cDNA samples was identical to that in clone 68616 but not either of the two duplications. These data argue that the only transcript present in brain corresponds to the intact locus, and that the two partial duplications are therefore not transcribed.

Immediately telomeric to *APBA2* lies another gene (*KIAA0574*; see Figure 2a), which was initially identified from a sequencing project of large, brain-derived cDNAs [25]. *KIAA0574* encodes a protein of unknown function, and is transcribed in the opposite orientation relative to *APBA2*. BLASTP using the *KIAA0574* predicted amino acid

Table 1: APBA2 intron-exon junctions

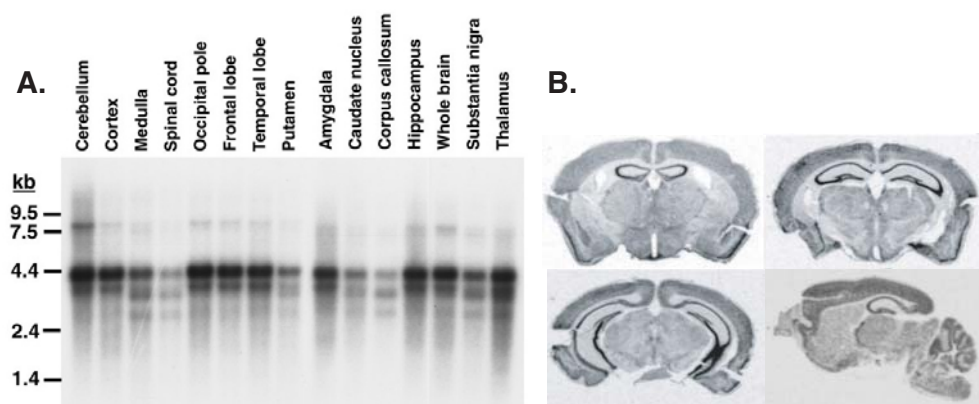
Exon	Exon Size (bp)	3' splice acceptor-5' splice donor*						Intron Size (kb)
1	117	TTT	ACA	...		TTC	TGA /gtgatg	>74
2	53	ttccag/ AAA	GAT			TTT	TAG /gtaaga	58
3	990	ccacag/ TGG	CTG	...		GAG	CAG /gtagga	20.3
						Glu	Gln	
4	80	ttccag/ GTT	TGC	...		CCA	GAG /gtaatt	1.1
		Val	Cys			Pro	Glu	
5	36	taacag/ ACA	AAG	...	GTG	GCT	G /gtaagt	17
		Thr	Lys		Val	Ala	Val	
6	145	cttcag/ TT	CCA	GGG	...	GTC	AAG /gtagag	1.1
		Val	Pro	Gly		Val	Lys	
7	35	tgtcag/ AGG	ATG	...		AAA	GCG /gtgtgt	4.2
		Arg	Met			Lys	Ala	
8	86	tccttag/ AAT	TCT	...		ACG	CAG /gtaagc	3
		Asn	Ser			Thr	Gln	
9	185	cgacag/ GAA	ACC	...		GAG	GAT /gtaagt	3.6
		Glu	Thr			Glu	Asp	
10	179	gcacag/ GCC	CAG	...		AAG	GAG /gtaagc	1
		Ala	Gln			Lys	Glu	
11	212	ttgcag/ CTG	CAG	...		ATC	AAG /gtaggc	1.5
		Leu	Gln			ISO	Lys	
12	119	ctgcag/ GGC	CTG	...		GGA	ATT /gtgagt	5.5
		Gly	Leu			Gly	ISO	
13	140	gcttag/ ATC	TGC	...		GGA	GAG /gtaagg	3
		ISO	Cys			Gly	Glu	
14	1257	ccgcag/ ATC	CAC	...				
		ISO	His					

*Intron and exon sequences are shown in lower-case and upper-case, respectively, with encoded amino acids indicated below the relevant codon.

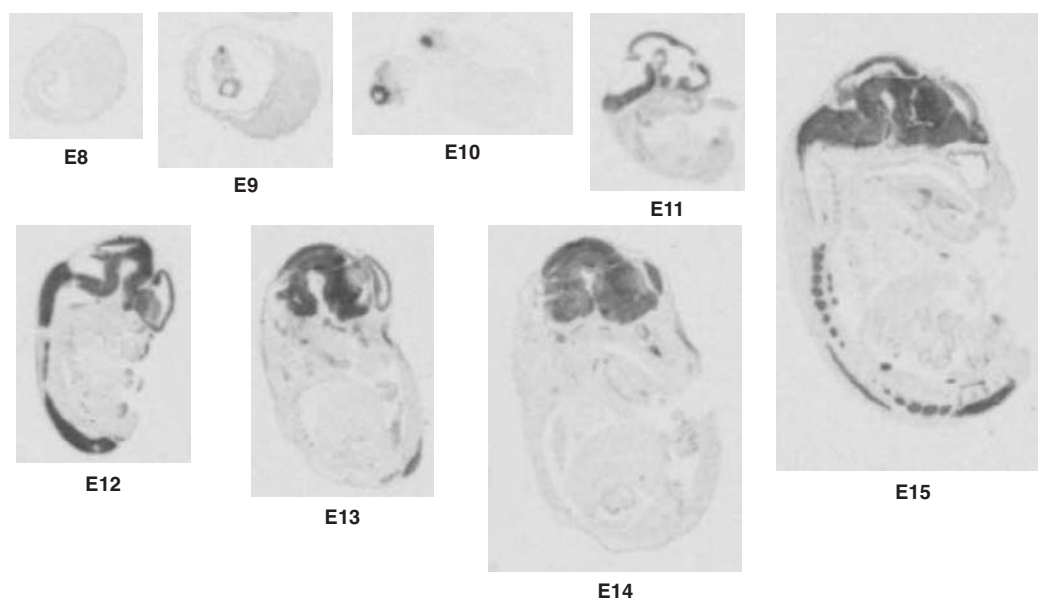
sequence detects weak homology (35% identity, 64% similarity over 69 residues) to a protein termed X123, the gene for which is located in the 9q12.21 Friedrich's ataxia region. The X123 gene in turn maps immediately adjacent to *APBA1* (alias Mint1 or X11 α). *APBA1* and *APBA2* are highly homologous for the C-terminal half of their respective sequences (90% similar, 84% identical). The N-terminal half, corresponding to the duplicated exon, is only weakly similar (36% similar, 30% identical). Thus the distinction between the *APBA1* and *APBA2* proteins corresponds to exon 3-encoded residues. Despite the comparatively weak homology between the KIAA0574 and X123 predicted peptides, this scenario suggests a clear evolutionary relationship between these gene pairs. The propensity of the 15q11-q13 region to undergo rearrangement may have played a role in the evolution of this gene family.

The distribution of *APBA2* brain expression was determined using *in situ* hybridization and northern blotting. Commercial brain northern blots were hybridized with a cDNA probe corresponding to exon 3. Such a probe allows us to discriminate between *APBA2* expression and

potential signal from *APBA1*. Northern blotting revealed a predominant moderate-abundance transcript of ~4.2-kb widely distributed throughout the brain and spinal cord (Figure 3a). Several smaller transcripts were also seen. A cDNA clone for the mouse ortholog of *APBA2* was obtained and a similar probe was used to test for expression using *in situ* hybridization to mouse brain sections. Figure 3b shows representative coronal and sagittal mouse brain sections. *Apba2* demonstrates moderate expression in mouse cortical and limbic regions including frontal, parietal, and temporal cortex, hippocampus, amygdala, thalamus, and cerebellum and lower level expression in many other regions summarized in Table 3. A developmental expression series illustrating expression from embryonic day 8 through day 15 (E8-E15) is shown in Figure 4. Early expression in the primitive neural tube emerges at day E10. *Apba2* expression extends throughout the neural tube during days E11 and E12, but is apparently restricted to the primitive brain vesicles during days E13 and E14. Within the brain, distribution is ubiquitous and uniform. In addition to diffuse expression throughout the brain, by day E15, a punctate pattern appears around the spinal

**Figure 3**

Expression analysis of APBA2/Apba2. A. Northern blot analysis of APBA2 in human brain. B. Relative expression of Mint2 in adult mouse brain by *in situ* hybridization. Mint2 shows highest expression in cortical and limbic regions, including limbic nuclei of the thalamus and hypothalamus. Brainstem, cerebellum, and occipital pole not done.

**Figure 4**

Developmental expression profile of Apba2. Apba2 expression in the primitive neural tube begins by embryonic day 10 (E10). Expression throughout the neural tube is seen in day E11 and E12, but appears to be restricted to the primitive brain vesicles by day E13 and continues during day E14 in all regions of primitive brain, including medulla, midbrain, thalamus, striatum, olfactory lobe, and neopallial cortex as seen above. By day E15, however, Apba2 transcript also appears in the dorsal root ganglia surrounding the spinal cord.

Table 2: PCR primer sequences

Exon	Exon Size (bp)		Primers	Ta (°C)	Product Size (bp)
1	117	F	5'-GAGCGTGTGGCCCATCGTT-3'	57	291
		R	5'-AGCACATCAGCCCAGACTCATTCA-3'		
2	53	F	5'-CATTTTTATACTGAAGGAGCTCT-3'	50	163
		R	5'-GCTGCCTTGATTTCTTACCTAA-3'		
3	991	F	5'-TGTGGGGTTTTGACTCTGTCCTTCC-3'	60	1098
		R	5'-GTACCCTGAGGTCCCTTGCTCTCG-3'		
4	83	F	5'-CTTTAACTTCCCCGTTTCCAGGTTTG-3'	57	137
		R	5'-TCCAGCCCAGAACTCTCATCCTTGAA-3'		
5	38	F	5'-GCATATTCAGGTTTTGCATCTTGC-3'	54	185
		R	5'-CAGCCCAGAACTCTCATCCTTGAA-3'		
6	146	F	5'-CCTGGCTGCCGTGTTCTTGTG-3'	61	316
		R	5'-CGCCATATTCCTCCTGCTAGTGC-3'		
7	34	F	5'-AGACCTGAGAGTGGGGGCATCAAC-3'	57	276
		R	5'-TCACCTACCAAGGGAAGAGGAAAAGC-3'		
8	88	F	5'-TCTGGTGATTAGGGATGTTGAGCA-3'	55	259
		R	5'-ATACCATGGGACAAGAAGGAGTCG-3'		
9	189	F	5'-CCACCATGCCAGCCCTGGAGTAC-3'	64	313
		R	5'-GCTGTGAACTTTGGGAGGGGAGTGC-3'		
10	179	F	5'-CCTGTGCCACGTGCCTGTCTCCAG-3'	62	253
		R	5'-CCTCCCTGAGGCTGGTGGGTGTGG-3'		
11	215	F	5'-TCCTCCTGTGGGTCTTGGCAGAG-3'	64	345
		R	5'-GGGGATGAAGTGAGGCAGGGTTGA-3'		
12	121	F	5'-TAAAGCCCAGCCTGTGACTCCTGTCC-3'	62	251
		R	5'-CCAGACATAGCCTGCCCCATTGGA-3'		
13	140	F	5'-CTGTCGGGTGTGGCGGAACACGT-3'	64	275
		R	5'-CTCTGGAGCGGGGAGGCTCATGG-3'		
14	1251	F	5'-GCGGCTCAGCCTCCTGTTTCTGTC-3'	62	200
		R	5'-CTGTGGCGTCGGGGTCAAGTCTGC-3'		

Forward (F) and reverse (R) oligonucleotide sequences are shown along with annealing temperatures (Ta) and PCR product size

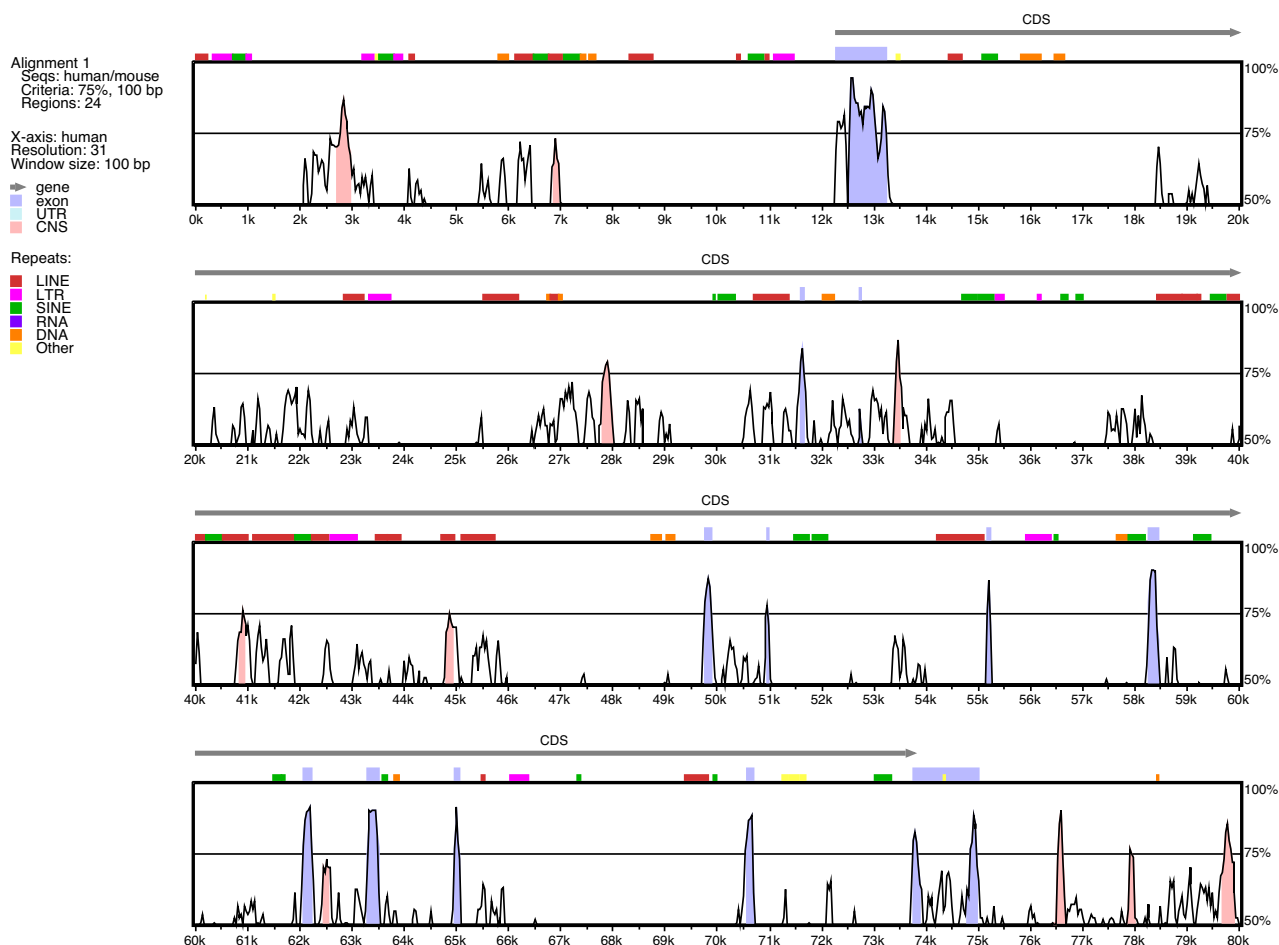
cord, corresponding to expression within the dorsal root ganglia.

The availability of emerging mouse genomic sequence allowed us to examine sequence conservation across the transcriptional unit. We submitted human and mouse genomic sequences to the VISTA web site <http://www.gsdlbl.gov/vista/> and homology is plotted and shown in Figure 5. Murine genomic sequence for exons 1 and 2 was not available, and the comparison was made for the remainder of the transcriptional unit. We would expect coding sequences to be conserved, but regions of non-coding conservation, which could harbour potential functional (conserved) sequences, are of particular interest. Several such regions are present within the *APBA2* transcriptional unit.

Discussion

We have described the mapping of *APBA2* to 15q13.1 within a region duplicated in idic(15) marker chromosomes, thus clarifying the map position of this gene. While public and private sequence assemblies now depict

the correct location for this gene, previous published reports indicated that this gene mapped proximal to BP3 within the PWS/AS deletion and autism-associated interstitial duplication interval [21,22]. STS content mapping with regional clones confirms the predicted genomic assemblies and places this gene distal to the BP3 *HERC2*-containing breakpoint. Two partial *APBA2* copies containing the first coding exon and downstream sequence and are located ~5-Mb telomeric in 15q13.3. This location contains duplcon-like sequences, based on BLAST-detected homology relationships to other non-overlapping BAC clones, which contain LCR15-inclusive sequences. This region may represent yet another functional duplcon. This suggests the possibility of such sequences upstream from *APBA2*, most likely corresponding to BP3 or in a sequence gap between *HERC2* and *APBA2*. The existence of a duplcon-like region near *APBA2* would most easily explain the duplication. The duplicated region corresponded to two blocks in the intact locus which are joined together in the duplication copies, perhaps reflecting the original organization of the intact ancestral locus. Within the 9.7 kb gap between homology blocks, another low-copy repeat se-

**Figure 5**

Comparative sequence analysis of human and mouse *APBA2* using VISTA. Output from VISTA analysis of the *ATPI0C* transcriptional unit is shown, with regions of non-coding sequence conservation (>75% identity) indicated by pink shading and coding homology by blue shading. Coding sequence (CDS) is indicated over the transcriptional unit with arrows indicating transcriptional orientation. Various repetitive sequences are indicated by shaded boxes over the homology plots.

quence of ~1.3 kb was identified. The pattern of partial duplication seen with *APBA2* is similar to that of *CHRNA7*. The repetitive nature of this region has presented significant challenges to genomic sequence assembly. It is likely that repeated elements cause some segments to be poorly represented in *E. coli*-based libraries. Our interest in this gene was driven by its plausibility as a candidate for autism, based upon function. The location of *APBA2* outside the narrow deletion/duplication interval makes this gene less attractive as a candidate for inherited autism. However, overexpression of this gene in *idic(15)* cases could be responsible in part for the increased phenotypic

severity in those cases compared to interstitial duplication.

We characterized gene structure for the human gene and expression for both human and mouse genes. There are 14 known exons, 13 of which are shown in the current public human genomic sequence assemblies. A putative first exon, corresponding to the 5' end of the Celera transcript, maps at least 74-kb upstream of exon 2. Based on a combination of the public and Celera assemblies, the *APBA2* transcriptional unit spans >197 kb. We provided primer sequences and amplification conditions to facilitate SNP and disease variant discovery at this locus.

Table 3: *Apba2* in situ hybridization summary

Structure	Expression*
Cortex	
Frontal (layers II-V)	++
Parietal (layers II-V)	++
Temporal (layers II-V)	++
Piriform (layer II)	+++
Cingulate	++
Insular	++
Thalamus	
Ant, PV, DM	++
Others	+
Hippocampus (CA1, CA2, CA3, DG)	+++
Amygdala	
Basolateral	++
Other	+
Septal Nuclei	
Lateral	++
Other	+
Habenula	
Medial	++
Lateral	+
Bed Nucleus of Stria Terminalis Hypothalamus	++
PO, Arc	++
Others	+

*Relative expression levels are indicated; +++ corresponds to the highest levels of expression and + to the lowest.

APBA2 is expressed throughout the central nervous system in the adult human and mouse and developing mouse embryos. The ~4.2-kb predominant transcript is larger than the longest available mRNA sequence of 3,647 bp. The difference is most likely due to as yet unidentified 5' untranslated exons or by use of alternative polyadenylation sites. The presence of multiple consensus polyadenylation signals both before and after the one utilized in this transcript is consistent with the latter hypothesis. Additionally, less abundant smaller transcripts exist. EST data supports the occurrence of alternative splicing for this gene. To our knowledge, transcriptional initiation site(s) have not been characterized for *APBA2*. The fact that the partial duplication copies are not expressed supports the premise that the intact locus may not have a transcriptional initiation site at exon 3.

Genomic sequence corresponding to the mouse ortholog is present in the public assembly, although sequence upstream of exon 3 is not available. VISTA comparison of human and mouse sequences reveals several sites of conserved non-coding sequences; these may contain functional elements necessary for proper expression.

Conclusions

The *APBA2* gene maps to 15q13.1 telomeric to *HERC2* and BP3. Partial duplications of a 6-kb segment containing the first coding exon are present as two copies at ~5-Mb telomeric in 15q13.3 and are not transcribed. The pattern of duplication for *APBA2* adds to a complex mosaic including many duplcon-like structures in 15q13. The tendency of this region to undergo rearrangement may bear relation to the evolution of this gene family.

Methods

PCR Analysis of *APBA2* Clones and Expression

PCR assays were performed with Roche *Taq* DNA polymerase (Indianapolis, IN) in 50 µL reactions for 30 cycles under standard conditions and T_a of 63°C, using 2.5 mM dNTPs and 0.5 µM primers. Primers: forward 5'-CAGGAGCCCCGAGAGCGAGGA-3'; reverse 5'-GGGTCTCCGGGGTGCTTGG-3'. Genomic and cDNA sources used as templates for amplification were obtained from commercial sources: RP-11 BACs from BACPAC Resources (Oakland, CA), YACs from the Whitehead Institute (WI) (Cambridge, MA), and Marathon cDNA from Invitrogen Life Technologies (Carlsbad, CA). All restriction digests were performed according to manufacturer recommendations using enzymes purchased from New England Biolabs (Beverly, MA).

Southern and Northern Blot Analyses

Non-conserved probes corresponding to exon 3 were hybridized to commercially-available Multiple Tissue Northern blots containing adult human brain polyA+ RNA (Invitrogen Life Technologies, Carlsbad, CA). Blots were hybridized overnight at 65°C according to manufacturer's recommendations and then washed for 2–3 hours in 2X SSC/0.1% SDS at 65°C. Filters were exposed to Kodak XAR x-ray film (Eastman Kodak; New Haven, CT).

In Situ Hybridization Analysis

Following perfusion, whole brains were obtained from C57BL/6 mice. Using a Leica CM300 cryostat (Leica, Wetzlar, Germany), frozen brains were sectioned into five series of 20 mm slices, with individual sections within a series separated by 100 µm. Wax-embedded embryonic sections were purchased from Novagen (Madison, WI). A probe-containing plasmid was developed for *Apba2*. Antisense and sense control cRNA probes were synthesized with the MAXIscript kit using T3, T7, or SP6 RNA polymerase (Ambion, Austin, TX) in reactions incorporating α-[³³P]-UTP (New England Nuclear Life Science Products, Boston, MA). Purified probes were hybridized to slides for 20 hours at 58°C and washed in a series of solutions to a final stringency of 0.1 X SSC at 65°C. Hybridized and washed slides were exposed to BioMax MR film (Eastman Kodak, New Haven, CT) for varying times,

depending on message abundance. X-ray images were prepared by scanning film at high resolution.

Sequence Analysis

Direct sequencing from BAC clone template was performed using ABI dye terminator chemistry according to manufacturer's recommendations (Applied Biosystems, Foster City, CA). Sequences were purified using spin columns and submitted to the Vanderbilt University Shared Sequencing Resource. Oligonucleotide primers were designed using Oligo v6.7 (Molecular Biology Insights, Inc., Cascade, CO) and MacVector v7.1 (Accelrys, Inc., Burlington, MA). BLAST comparisons were performed from the NCBI <http://www.ncbi.nlm.nih.gov> or Ensembl <http://www.ensembl.org/> websites. Comparative analysis of human and mouse genomic sequences was performed through the VISTA web site <http://www.gsd.lbl.gov/vista/>.

List of abbreviations

PWS, Prader-Willi syndrome

AS, Angelman syndrome

idic(15), pseudodicentric, inverted, duplicated chromosomes 15

BP, breakpoint

LCR15, low copy repeats on chromosome 15

STS, sequence-tagged site

EST, expressed sequence tag

YAC, yeast artificial chromosome

BAC, bacterial artificial chromosome

PCR, polymerase chain reaction

RFLP, restriction fragment length polymorphism

Authors' contributions

JSS and ELN conceived of the study and participated in its design and coordination. JSS drafted the manuscript. MKH isolated and mapped the KIAA0574 gene and participated in the characterization of APBA2. TA performed VISTA comparison and multiple other DNA sequence analyses. ELN designed, developed, and executed all PCR-based experiments and most *in silico* DNA sequence analysis, and performed all hybridization assays.

Acknowledgements

This work was supported by grant MH61009 to JSS.

References

1. Ji Y, Eichler EE, Schwartz S and Nicholls RD **Structure of chromosomal duplicons and their role in mediating human genomic disorders** *Genome Res.* 2000, **10**:597-610
2. Ledbetter DH and Ballabio A **Molecular cytogenetics of contiguous gene syndromes: mechanisms and consequences of gene dosage imbalance** In: *The Metabolic and Molecular Basis of Inherited Disease* (Edited by: Scriver CR, Beaudet AL, Sly WS, Valle D) New York: McGraw Hill 1995, 811-839
3. Nicholls RD and Knepper JL **Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes** *Ann Rev Genomics Hum Genet* 2001, **2**:153-75
4. Jiang YH, Tsai TF, Bressler J and Beaudet AL **Imprinting in Angelman and Prader-Willi Syndromes** *Curr Opin Genet Dev* 1998, **8**:334-342
5. Cook EH Jr, Lindgren V, Leventhal BL, Courchesne R, Lincoln A, Shulman C, Lord C and Courchesne E **Autism or atypical autism in maternally but not paternally derived proximal 15q duplication** *Am J Hum Genet* 1997, **60**:928-34
6. Bolton PF, Dennis NR, Browne CE, Thomas NS, Veltman MW, Thompson RJ and Jacobs P **The phenotypic manifestations of interstitial duplications of proximal 15q with special reference to the autistic spectrum disorders** *Am J Med Genet* 2001, **105**:675-85
7. Roberts SE, Dennis NR, Browne CE, Willatt L, Woods G, Cross I, Jacobs PA and Thomas S **Characterisation of interstitial duplications and triplications of chromosome 15q11-q13** *Hum Genet* 2002, **110**:227-34
8. Huang B, Crolla JA, Christian SL, Wolf-Ledbetter ME, Macha ME, Papenhausen PN and Ledbetter DH **Refined Molecular Characterization Of the Breakpoints In Small Inv Dup(15) Chromosomes** *Hum Genet* 1997, **99**:11-17
9. Robinson WP, Binkert F, Gine R, Vazquez C, Muller W, Rosenkranz W and Schinzel A **Clinical and molecular analysis of five inv dup(15) patients** *Eur J Hum Genet* 1993, **1**:37-50
10. Clayton-Smith J, Webb T, Cheng XJ, Pembrey ME and Malcolm S **Duplication of chromosome 15 in the region 15q11-13 in a patient with developmental delay and ataxia with similarities to Angelman syndrome** *J Med Genet* 1993, **30**:529-31
11. Flejter WL, Bennett-Baker PE, Ghaziuddin M, McDonald M, Sheldon S and Gorski JL **Cytogenetic and molecular analysis of inv dup(15) chromosomes observed in two patients with autistic disorder and mental retardation** *Am J Med Genet* 1996, **61**:182-7
12. Schinzel AA, Brecevic L, Bernasconi F, Binkert F, Berthet F, Wuilloud A and Robinson WP **Intrachromosomal triplication of 15q11-q13** *J Med Genet* 1994, **31**:798-803
13. Ungaro P, Christian SL, Fantes JA, Mutirangura A, Black S, Reynolds J, Malcolm S, Dobyns WB and Ledbetter DH **Molecular characterisation of four cases of intrachromosomal triplication of chromosome 15q11-q14** *J Med Genet* 2001, **38**:26-34
14. Christian SL, Robinson WP, Huang B, Mutirangura A, Line MR, Nakao M, Surti U, Chakravarti A and Ledbetter DH **Molecular characterization of two proximal deletion breakpoint regions in both Prader-Willi and Angelman syndrome patients** *Am J Hum Genet* 1995, **57**:40-8
15. Christian SL, Fantes JA, Mewborn SK, Huang B and Ledbetter DH **Large genomic duplicons map to sites of instability in the Prader-Willi/Angelman syndrome chromosome region (15q11-q13)** *Hum Mol Genet* 1999, **8**:1025-1037
16. Amos-Landgraf JM, Ji YG, Gottlieb W, Depinet T, Wandstrat AE, Cassidy SB, Driscoll DJ, Rogan PK, Schwartz S and Nicholls RD **Chromosome breakage in the Prader-Willi and Angelman syndromes involves recombination between large, transcribed repeats at proximal and distal breakpoints** *Am J Hum Genet* 1999, **65**:370-386
17. Ji YG, Walkowicz MJ, Buiting K, Johnson DK, Tarvin RE, Rinchik EM, Horsthemke B, Stubbs L and Nicholls RD **The ancestral gene for transcribed, low-copy repeats in the Prader-Willi/Angelman region encodes a large protein implicated in protein trafficking, which is deficient in mice with neuromuscular and spermiogenic abnormalities** *Hum Mol Genet* 1999, **8**:533-542
18. Pujana MA, Nadal M, Gratacos M, Peral B, Csiszar K, Gonzalez-Sarmiento R, Sumoy L and Estivill X **Additional complexity on human chromosome 15q: identification of a set of newly recognized duplicons (LCR15) on 15q11-q13, 15q24, and 15q26** *Genome Res* 2001, **11**:98-111

19. Pujana MA, Nadal M, Guitart M, Armengol L, Gratacos M and Estivill X **Human chromosome 15q11-q14 regions of rearrangements contain clusters of LCR15 duplicons** *Eur J Hum Genet* 2002, **10**:26-35
20. Blanco G, Irving NG, Brown SD, Miller CC and McLoughlin DM **Mapping of the human and murine XH-like genes (APBA2 and apba2), the murine Fe65 gene (Apbb1), and the human Fe65-like gene (APBB2): genes encoding phosphotyrosine-binding domain proteins that interact with the Alzheimer's disease amyloid precursor protein** *Mamm Genome* 1998, **9**:473-5
21. Maddox LO, Menold MM, Bass MP, Rogala AR, Pericak-Vance MA, Vance JM and Gilbert JR **Autistic disorder and chromosome 15q11-q13: construction and analysis of a BAC/PAC contig** *Genomics* 1999, **62**:325-31
22. Christian SL, Bhatt NK, Martin SZ, Sutcliffe JS, Kubota T, Huang B, Mutirangura A, Chinault AC, Beaudet AL and Ledbetter DH **Integrated Yac Contig Map of the Prader-Willi/Angelman Region On Chromosome 15q11-Q13 With Average Sts Spacing of 35 Kb** *Genome Res* 1998, **8**:146-157
23. Biederer T and Sudhof TC **Mints as adaptors. Direct binding to neurexins and recruitment of munc 8** *J Biol Chem* 2000, **275**:39803-6
24. Gault J, Robinson M, Berger R, Drebing C, Logel J, Hopkins J, Moore T, Jacobs S, Meriwether J and Choi MJ **Genomic organization and partial duplication of the human alpha7 neuronal nicotinic acetylcholine receptor gene (CHRNA7)** *Genomics* 1998, **52**:173-85
25. Nagase T, Ishikawa K, Miyajima N, Tanaka A, Kotani H, Nomura N and Ohara O **Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro** *DNA Res* 1998, **5**:31-9

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

